

**Exhibit 13 to Complaint**  
**Intellectual Ventures I LLC and Intellectual Ventures II LLC**

**Example Southwest Count 6 Systems and Services**  
**U.S. Patent No. 7,257,582 (“’582 Patent”)**

The Accused Systems and Services include without limitation Southwest systems and services that utilize Hadoop; all past, current, and future systems and services that operate in the same or substantially similar manner as the specifically identified systems and services; and all past, current, and future Southwest systems and services that have the same or substantially similar features as the specifically identified systems and services (“Example Southwest Count 6 Systems and Services” or “Southwest Systems and Services”).<sup>1</sup>

On information and belief, the Southwest Systems and Services use Hadoop in its private cloud(s). For example, Southwest posts, or has posted, job opportunities that require familiarity with Hadoop concepts.

*See <https://www.linkedin.com/in/meghanaa-malleboina/>,* job profile of data engineer who stated that they enhanced Hadoop YARN cluster efficiency by 35% through optimized resource allocation.

*See <https://www.linkedin.com/in/rohan-kumar-9861bb127>,* job profile of Senior Hadoop Developer who stated that they are/were a Hadoop developer for Southwest Airlines. (last accessed 10/8/24).

*See <https://www.linkedin.com/in/harish-t-b82094128>,* job profile of Senior Hadoop Developer who stated that they are/were a Hadoop developer for Southwest Airlines. (last accessed 10/8/24).

As another example, Southwest has stated that it is investing in cloud technology and has “moved about 50% of its technology” to the cloud and has indicated cloud migration is one of its areas of focus for 2024 and beyond. Source:

<https://www.phocuswire.com/southwest-airlines-cio-tech-investment>.

---

<sup>1</sup> For the avoidance of doubt, Plaintiffs do not accuse public clouds of Southwest if those services are provided by a cloud provider with a license to Plaintiffs’ patents that covers Southwest’s activities. IV will provide relevant license agreements for cloud providers in discovery. To the extent any of these licenses are relevant to Southwest’s activities, Plaintiffs will meet and confer with Southwest about the impact of such license(s).

On information and belief, other information confirms Southwest uses Hadoop technology.

## How Data Analytics Make Airlines Fly

Look up at the sky and think in numbers: Each day, nearly 7,000 commercial aircraft take off on 24,000 flights, according to [the Federal Aviation Administration](#). With the expert guidance of 14,000 air traffic controllers, they fly about 2.2 million passengers every 24 hours. To do that, airlines and controllers coordinate airplanes' movements through internal dispatch offices and 476 control towers. Add to the mix some 200,000 general aviation aircraft traveling between more than 19,000 airports, and you can envision the logistical challenges involved in moving [719 million passengers](#) around the U.S. each year. It doesn't take much to imagine the complex logistics involved in making a system of such scale work properly. While "data analytics" is a relatively new term to the mainstream, airlines have been applying such principles to the business for years, often under the phrase "operations modeling," said [Doug Gray, director of enterprise data analytics for Dallas-based Southwest Airlines](#). At first, the data was used to guide decisions about fueling, crew schedules and flight itineraries. Today, analytics are used across the organization, in functions from marketing to operations, explained Gray, who's also a member of the International Institute for Analytics' Expert Network. "They have a significant impact on costs and profitability," he said. Indeed, calling real-time analytics "integral" to any airline's success isn't going too far. While specialists in fueling, crew scheduling, flight scheduling and other areas may have their plans laid out perfectly, their work is never immune from unexpected developments. "Daily operations is where stuff hits the fan," Gray observed. "You can't predict a heart attack in-flight, or a control-tower fire." Schedulers have only so much advance warning of bad weather. Even on the best of days, it seems an airline's plans can only be regarded as tentative.

## Industry Expertise Helps

Of course, airlines need people who can implement and maintain analytics systems—software engineers and developers who need to have at least some expertise with data, Gray said. For his part, Gray is particularly interested in tech pros familiar with Oracle, Teradata and Amazon Cloud Services. Also important are relational database skills such as NoSQL and Mongo DB. Although he sees the company experimenting more with Hadoop (“It’s better in the cloud.”), Gray said that “unless something better comes along,” the company will continue to rely heavily on R and Alteryx, a “data science self-service desktop” that combines ETL, R and visualization in one GUI-driven application. Like many employers, Southwest will hire “the right person and train them” on needed skills, especially if they’re just out of school, Gray said. And, he believes, industry experience can give candidates an advantage. As more people pursue careers using data science, analytics and operations research, “more departments will have their own [data and analytics] experts, joined together by a center of excellence.” “Our biggest constraint isn’t data,” he added. “It’s teaming up the right data people with the right subject-matter experts. There’s going to be a battle for talent, and we need people with a passion for the airline business.”

Source: <https://www.dice.com/career-advice/how-data-analytics-airlines-fly>.<sup>2</sup>

---

<sup>2</sup> All sources cited in this document were publicly accessible as of the filing date of the Complaint.

U.S. Patent No. 7,257,582 (Claim 1)	
Claim(s)	Example Southwest Count 6 Systems and Services
1. A method of effecting on a preexisting input file a computer-executable process comprised of a plurality of subtasks, the method comprising the steps of:	<p>To the extent this preamble is limiting, on information and belief, the Southwest Count 6 Systems and Services practice a method of effecting on a preexisting input file a computer-executable process comprised of a plurality of subtasks.</p> <p>On information and belief, Hadoop includes MapReduce, a software framework that processes multiple map tasks in parallel. An input data set is split into multiple chunks and each chunk is processed by a map task.</p> <p><b>Overview</b></p> <p>Hadoop MapReduce is a software framework for easily writing applications which process vast amounts of data (multi-terabyte data-sets) in-parallel on large clusters (thousands of nodes) of commodity hardware in a reliable, fault-tolerant manner.</p> <p>A MapReduce job usually splits the input data-set into independent chunks which are processed by the <i>map tasks</i> in a completely parallel manner. The framework sorts the outputs of the maps, which are then input to the <i>reduce tasks</i>. Typically both the input and the output of the job are stored in a file-system. The framework takes care of scheduling tasks, monitoring them and re-executes the failed tasks.</p> <p>Typically the compute nodes and the storage nodes are the same, that is, the MapReduce framework and the Hadoop Distributed File System (see <a href="#">HDFS Architecture Guide</a>) are running on the same set of nodes. This configuration allows the framework to effectively schedule tasks on the nodes where data is already present, resulting in very high aggregate bandwidth across the cluster.</p> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p>

U.S. Patent No. 7,257,582 (Claim 1)	
Claim(s)	Example Southwest Count 6 Systems and Services
	<p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p>
(a) automatically determining file allocation and logically subdividing records of said input file into a plurality of partitions;	<p>On information and belief, the Southwest Count 6 Systems and Services practice automatically determining file allocation and logically subdividing records of said input file into a plurality of partitions.</p> <p>On information and belief, Hadoop's InputFormat describes the input-specification for a MapReduce job. The input file is logically split into multiple InputSplit instances.</p>

U.S. Patent No. 7,257,582 (Claim 1)	
Claim(s)	Example Southwest Count 6 Systems and Services
	<p><b>Job Input</b></p> <p>InputFormat describes the input-specification for a MapReduce job.</p> <p>The MapReduce framework relies on the InputFormat of the job to:</p> <ol style="list-style-type: none"> <li>1. Validate the input-specification of the job.</li> <li>2. Split-up the input file(s) into logical InputSplit instances, each of which is then assigned to an individual Mapper.</li> <li>3. Provide the RecordReader implementation used to glean input records from the logical InputSplit for processing by the Mapper.</li> </ol> <p>The default behavior of file-based InputFormat implementations, typically sub-classes of FileInputFormat, is to split the input into <i>logical</i> InputSplit instances based on the total size, in bytes, of the input files. However, the FileSystem blocksize of the input files is treated as an upper bound for input splits. A lower bound on the split size can be set via mapreduce.input.fileinputformat.split.minsize.</p> <p>Clearly, logical splits based on input-size is insufficient for many applications since record boundaries must be respected. In such cases, the application should implement a RecordReader, who is responsible for respecting record-boundaries and presents a record-oriented view of the logical InputSplit to the individual task.</p> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p>
(b) distributing descriptions of all of said partitions to each of a plurality of subtask processors	<p>On information and belief, the Southwest Count 6 Systems and Services practice distributing descriptions of all of said partitions to each of a plurality of subtask processors.</p> <p>On information and belief, Hadoop's InputSplit represents data that presents a byte oriented view of the input. A recordreader converts the byte oriented view of the data to a record oriented view and presents it to a mapper. Mappers run inside a compute node.</p> <p>Typically the compute nodes and the storage nodes are the same, that is, the MapReduce framework and the Hadoop Distributed File System (see HDFS Architecture Guide) are running on the same set of nodes. This configuration allows the framework to effectively schedule tasks on the nodes where data is already present, resulting in very high aggregate bandwidth across the cluster.</p> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p>

U.S. Patent No. 7,257,582 (Claim 1)	
Claim(s)	Example Southwest Count 6 Systems and Services
	<p><b>Job Input</b></p> <p><code>InputFormat</code> describes the input-specification for a MapReduce job.</p> <p>The MapReduce framework relies on the <code>InputFormat</code> of the job to:</p> <ol style="list-style-type: none"> <li>1. Validate the input-specification of the job.</li> <li>2. Split-up the input file(s) into logical <code>InputSplit</code> instances, each of which is then assigned to an individual <code>Mapper</code>.</li> <li>3. Provide the <code>RecordReader</code> implementation used to glean input records from the logical <code>InputSplit</code> for processing by the <code>Mapper</code>.</li> </ol> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p> <p>If <code>TextInputFormat</code> is the <code>InputFormat</code> for a given job, the framework detects input-files with the <code>.gz</code> extensions and automatically decompresses them using the appropriate <code>CompressionCodec</code>. However, it must be noted that compressed files with the above extensions cannot be <i>split</i> and <code>each compressed file is processed in its entirety by a single mapper</code>.</p> <p><b>InputSplit</b></p> <p><code>InputSplit</code> represents the data to be processed by an individual <code>Mapper</code>.</p> <p>Typically <code>InputSplit</code> presents a byte-oriented view of the input, and it is the responsibility of <code>RecordReader</code> to process and present a record-oriented view.</p> <p><code>FileSplit</code> is the default <code>InputSplit</code>. It sets <code>mapreduce.map.input.file</code> to the path of the input file for the logical split.</p> <p><b>RecordReader</b></p> <p><code>RecordReader</code> reads <code>&lt;key, value&gt;</code> pairs from an <code>InputSplit</code>.</p> <p>Typically the <code>RecordReader</code> converts the byte-oriented view of the input, provided by the <code>InputSplit</code>, and presents a record-oriented to the <code>Mapper</code> implementations for processing. <code>RecordReader</code> thus assumes the responsibility of processing record boundaries and presents the tasks with keys and values.</p> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p>

U.S. Patent No. 7,257,582 (Claim 1)	
Claim(s)	Example Southwest Count 6 Systems and Services
HDFS Architecture	
	<p>The diagram illustrates the HDFS architecture. At the top, a <b>Namenode</b> (blue rectangle) contains a box labeled "Metadata (Name, replicas, ...): /home/foo/data, 3, ...". A <b>Client</b> (orange oval) sends "Metadata ops" to the Namenode. The Namenode also sends "Block ops" to Datanodes. Below the Namenode, a red box encloses two racks. <b>Rack 1</b> contains two Datanodes, each with four green squares representing blocks. One Datanode in Rack 1 sends "Read" operations to the Client. The other Datanode in Rack 1 sends "Replication" operations to a Datanode in <b>Rack 2</b>. <b>Rack 2</b> contains two Datanodes, each with four green squares. One Datanode in Rack 2 contains a box labeled "Blocks". A <b>Client</b> (orange oval) sends "Write" operations to the Datanode in Rack 2.</p> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-project-dist/hadoop-hdfs/HdfsDesign.html">https://hadoop.apache.org/docs/stable/hadoop-project-dist/hadoop-hdfs/HdfsDesign.html</a>.</p>

(c) simultaneously executing at least a respective one of the subtasks of the computer-executable process in each of at least some of said processors on a respective one of the partitions with each subtask reading and processing the respective

On information and belief, the Southwest Count 6 Systems and Services practice simultaneously executing at least a respective one of the subtasks of the computer-executable process in each of at least some of said processors on a respective one of the partitions with each subtask reading and processing the respective partition so as to process the respective partition and produce respective subtask output.  
On information and belief, Multiple map tasks are executed in parallel. Each Maptask processes a respective InputSplit.

U.S. Patent No. 7,257,582 (Claim 1)	
Claim(s)	Example Southwest Count 6 Systems and Services
partition so as to process the respective partition and produce respective subtask output and;	<p><b>Overview</b></p> <p>Hadoop MapReduce is a software framework for easily writing applications which process vast amounts of data (multi-terabyte data-sets) in-parallel on large clusters (thousands of nodes) of commodity hardware in a reliable, fault-tolerant manner.</p> <p>A MapReduce job usually splits the input data-set into independent chunks which are processed by the map tasks in a completely parallel manner. The framework sorts the outputs of the maps, which are then input to the reduce tasks. Typically both the input and the output of the job are stored in a file-system. The framework takes care of scheduling tasks, monitoring them and re-executes the failed tasks.</p> <p>Typically the compute nodes and the storage nodes are the same, that is, the MapReduce framework and the Hadoop Distributed File System (see <a href="#">HDFS Architecture Guide</a>) are running on the same set of nodes. This configuration allows the framework to effectively schedule tasks on the nodes where data is already present, resulting in very high aggregate bandwidth across the cluster.</p> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p> <p><b>Job Input</b></p> <p><code>InputFormat</code> describes the input-specification for a MapReduce job.</p> <p>The MapReduce framework relies on the <code>InputFormat</code> of the job to:</p> <ol style="list-style-type: none"> <li>1. Validate the input-specification of the job.</li> <li>2. Split-up the input file(s) into logical <code>InputSplit</code> instances, each of which is then assigned to an individual <code>Mapper</code>.</li> <li>3. Provide the <code>RecordReader</code> implementation used to glean input records from the logical <code>InputSplit</code> for processing by the <code>Mapper</code>.</li> </ol> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p>

U.S. Patent No. 7,257,582 (Claim 1)	
Claim(s)	Example Southwest Count 6 Systems and Services
	<p><b>Mapper</b></p> <p>Mapper maps input key/value pairs to a set of intermediate key/value pairs.</p> <p>Maps are the individual tasks that transform input records into intermediate records. The transformed intermediate records do not need to be of the same type as the input records. A given input pair may map to zero or many output pairs.</p> <p>The Hadoop MapReduce framework spawns one map task for each InputSplit generated by the InputFormat for the job.</p> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p> <p><b>How Many Maps?</b></p> <p>The number of maps is usually driven by the total size of the inputs, that is, the total number of blocks of the input files.</p> <p>The right level of parallelism for maps seems to be around 10-100 maps per-node, although it has been set up to 300 maps for very cpu-light map tasks. Task setup takes a while, so it is best if the maps take at least a minute to execute.</p> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p> <p>On information and belief, Each MapTask generates an intermediate map output.</p> <p><b>Overview</b></p> <p>Hadoop MapReduce is a software framework for easily writing applications which process vast amounts of data (multi-terabyte data-sets) in-parallel on large clusters (thousands of nodes) of commodity hardware in a reliable, fault-tolerant manner.</p> <p>A MapReduce job usually splits the input data-set into independent chunks which are processed by the map tasks in a completely parallel manner. The framework sorts the outputs of the maps, which are then input to the reduce tasks. Typically both the input and the output of the job are stored in a file-system. The framework takes care of scheduling tasks, monitoring them and re-executes the failed tasks.</p> <p>Typically the compute nodes and the storage nodes are the same, that is, the MapReduce framework and the Hadoop Distributed File System (see <a href="#">HDFS Architecture Guide</a>) are running on the same set of nodes. This configuration allows the framework to effectively schedule tasks on the nodes where data is already present, resulting in very high aggregate bandwidth across the cluster.</p>

U.S. Patent No. 7,257,582 (Claim 1)	
Claim(s)	Example Southwest Count 6 Systems and Services
	<p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p> <p><b>InputSplit</b></p> <p>InputSplit represents the data to be processed by an individual Mapper.</p> <p>Typically InputSplit presents a byte-oriented view of the input, and it is the responsibility of RecordReader to process and present a record-oriented view.</p> <p>FileSplit is the default InputSplit. It sets mapreduce.map.input.file to the path of the input file for the logical split.</p> <p><b>RecordReader</b></p> <p>RecordReader reads &lt;key, value&gt; pairs from an InputSplit.</p> <p>Typically the RecordReader converts the byte-oriented view of the input, provided by the InputSplit, and presents a record-oriented to the Mapper implementations for processing. RecordReader thus assumes the responsibility of processing record boundaries and presents the tasks with keys and values.</p> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p> <p><b>Mapper</b></p> <p>Mapper maps input key/value pairs to a set of intermediate key/value pairs.</p> <p>Maps are the individual tasks that transform input records into intermediate records. The transformed intermediate records do not need to be of the same type as the input records. A given input pair may map to zero or many output pairs.</p> <p>The Hadoop MapReduce framework spawns one map task for each InputSplit generated by the InputFormat for the job.</p> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p>

U.S. Patent No. 7,257,582 (Claim 1)	
Claim(s)	Example Southwest Count 6 Systems and Services
	<p><b>Data Compression</b></p> <p>Hadoop MapReduce provides facilities for the application-writer to specify compression for both intermediate map-outputs and the job-outputs i.e. output of the reduces. It also comes bundled with <code>CompressionCodec</code> implementation for the <code>zlib</code> compression algorithm. The <code>gzip</code>, <code>bzip2</code>, <code>snappy</code>, and <code>lz4</code> file format are also supported.</p> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p>
(d) thereafter repeating step (c) in at least some of the subtask processors each with another unprocessed partition on a first-come/first-served basis; and	<p>On information and belief, the Southwest Count 6 Systems and Services practice repeating step (c) in at least some of the subtask processors each with another unprocessed partition on a first-come/first-served basis.</p> <p>On information and belief, Hadoop includes a FIFO scheduler. Each mapper processes an InputSplit at a time. For example, a mapper processes one line of the word count application at a time.</p> <p><b>Overview</b></p> <p><b>Overview</b></p> <p>The Yarn scheduler is a fertile area of interest with different implementations, e.g., Fifo, Capacity and Fair schedulers. Meanwhile, several optimizations are also made to improve scheduler performance for different scenarios and workload. Each scheduler algorithm has its own set of features, and drives scheduling decisions by many factors, such as fairness, capacity guarantee, resource availability, etc. It is very important to evaluate a scheduler algorithm very well before we deploy in a production cluster. Unfortunately, currently it is non-trivial to evaluate a scheduler algorithm. Evaluating in a real cluster is always time and cost consuming, and it is also very hard to find a large-enough cluster. Hence, a simulator which can predict how well a scheduler algorithm for some specific workload would be quite useful.</p> <p>Source: <a href="https://hadoop.apache.org/docs/r2.9.2/hadoop-sls/SchedulerLoadSimulator.html">https://hadoop.apache.org/docs/r2.9.2/hadoop-sls/SchedulerLoadSimulator.html</a>.</p> <p><b>Example: WordCount v1.0</b></p> <p>Before we jump into the details, lets walk through an example MapReduce application to get a flavour for how they work. WordCount is a simple application that counts the number of occurrences of each word in a given input set.</p>

U.S. Patent No. 7,257,582 (Claim 1)	
Claim(s)	Example Southwest Count 6 Systems and Services
	<p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p> <p> <b>Walk-through</b></p> <p>The WordCount application is quite straight-forward.</p> <pre>public void map(Object key, Text value, Context context                ) throws IOException, InterruptedException {     StringTokenizer itr = new StringTokenizer(value.toString());     while (itr.hasMoreTokens()) {         word.set(itr.nextToken());         context.write(word, one);     } }</pre> <p>The Mapper implementation, via the <code>map</code> method, processes one line at a time, as provided by the specified <code>TextInputFormat</code>. It then splits the line into tokens separated by whitespaces, via the <code>StringTokenizer</code>, and emits a key-value pair of &lt; <code>&lt;word&gt;</code>, 1&gt;.</p> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p>

U.S. Patent No. 7,257,582 (Claim 1)	
Claim(s)	Example Southwest Count 6 Systems and Services
	<p>For the given sample input the first map emits:</p> <div style="border: 1px solid black; padding: 10px;"><pre>&lt; Hello, 1&gt; &lt; World, 1&gt; &lt; Bye, 1&gt; &lt; World, 1&gt;</pre></div> <p>The second map emits:</p> <div style="border: 1px solid black; padding: 10px;"><pre>&lt; Hello, 1&gt; &lt; Hadoop, 1&gt; &lt; Goodbye, 1&gt; &lt; Hadoop, 1&gt;</pre></div> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p>

U.S. Patent No. 7,257,582 (Claim 1)	
Claim(s)	Example Southwest Count 6 Systems and Services
	<p>WordCount also specifies a combiner. Hence, the output of each map is passed through the local combiner (which is same as the Reducer as per the job configuration) for local aggregation, after being sorted on the keys.</p> <p>The output of the first map:</p> <pre>&lt; Bye, 1&gt; &lt; Hello, 1&gt; &lt; World, 2&gt;</pre> <p>The output of the second map:</p> <pre>&lt; Goodbye, 1&gt; &lt; Hadoop, 2&gt; &lt; Hello, 1&gt;</pre> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p>
(e) generating at least one output combining all of the subtask outputs and reflecting the processing of all of said subtasks.	<p>On information and belief, the Southwest Count 6 Systems and Services practice generating at least one output combining all of the subtask outputs and reflecting the processing of all of said subtasks.</p> <p>On information and belief, intermediate values are grouped by the Hadoop framework and passed to a Reducer to determine the final output.</p>

U.S. Patent No. 7,257,582 (Claim 1)	
Claim(s)	Example Southwest Count 6 Systems and Services
	<p><b>Mapper</b></p> <p>Mapper maps input key/value pairs to a set of intermediate key/value pairs.</p> <p>Maps are the individual tasks that transform input records into intermediate records. The transformed intermediate records do not need to be of the same type as the input records. A given input pair may map to zero or many output pairs.</p> <p>The Hadoop MapReduce framework spawns one map task for each InputSplit generated by the InputFormat for the job.</p> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p> <p>All intermediate values associated with a given output key are subsequently grouped by the framework, and passed to the Reducer(s) to determine the final output. Users can control the grouping by specifying a Comparator via Job.setGroupingComparatorClass(Class).</p> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p> <p>The Mapper outputs are sorted and then partitioned per Reducer. The total number of partitions is the same as the number of reduce tasks for the job. Users can control which keys (and hence records) go to which Reducer by implementing a custom Partitioner.</p> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p> <p><b>Data Compression</b></p> <p>Hadoop MapReduce provides facilities for the application-writer to specify compression for both intermediate map-outputs and the job-outputs i.e. output of the reduces. It also comes bundled with <code>CompressionCodec</code> implementation for the <code>zlib</code> compression algorithm. The <code>gzip</code>, <code>bzip2</code>, <code>snappy</code>, and <code>Iz4</code> file format are also supported.</p> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p> <p>On information and belief, the Reducer reduces a set of intermediate values to a smaller output.</p>

U.S. Patent No. 7,257,582 (Claim 1)	
Claim(s)	Example Southwest Count 6 Systems and Services
	<p><b>Reducer</b></p> <p>Reducer reduces a set of intermediate values which share a key to a smaller set of values.</p> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p> <p>The Reducer implementation, via the reduce method just sums up the values, which are the occurrence counts for each key (i.e. words in this example).</p> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p> <p>WordCount also specifies a combiner. Hence, the output of each map is passed through the local combiner (which is same as the Reducer as per the job configuration) for local aggregation, after being sorted on the keys.</p> <p>The output of the first map:</p> <div style="border: 1px solid black; padding: 10px;"> <pre>&lt; Bye, 1&gt; &lt; Hello, 1&gt; &lt; World, 2&gt;</pre> </div> <p>The output of the second map:</p> <div style="border: 1px solid black; padding: 10px;"> <pre>&lt; Goodbye, 1&gt; &lt; Hadoop, 2&gt; &lt; Hello, 1&gt;</pre> </div> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a></p>

U.S. Patent No. 7,257,582 (Claim 1)	
Claim(s)	Example Southwest Count 6 Systems and Services
	<p><a href="#">core/MapReduceTutorial.html</a>.</p> <p>Thus the output of the job is:</p> <div style="border: 1px solid black; padding: 10px;"><pre>&lt; Bye, 1&gt; &lt; Goodbye, 1&gt; &lt; Hadoop, 2&gt; &lt; Hello, 2&gt; &lt; World, 2&gt;</pre></div> <p>Source: <a href="https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html">https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html</a>.</p>